Towards the Development of a Clinical Safety Acceptance Rule for Low-Field MRI using Hybridized Uncertainty-Aware Autoencoders

Mberekpe Emmanuel Funperefagha[1], Nafisa Opemi Jimoh[2], Usman Tijjani[3], Oyewole Oyelami[4], Joshua Taiwo Okegbemi[5], Hafsat Muhammed Usman[3], Olumide Olayinka Obe[5], William Thisleton[6], Confidence Raymond, Aondona Moses Iorumbur[7], Udunna Anazodo[8], Maruf Adewole[7], Adegboyega Adebayo[6,9]

[1] Federal University of Lafia, Nasarawa, Nigeria
[2]Department of Computer Engineering, Veritas University, Abuja, Nigeria
[3]Department of Medical Radiography, Federal University of Lafia
[4]Department of Mathematics, National Open University of Nigeria, Abuja, Nigeria
[5]Computer Science Department, Federal University of Technology, Akure, Nigeria
[6]AIX Lab, State University of New York, Polytechnic Institute, Utica, New York, USA
[7]Medical Artificial Intelligence Lab, Lagos, Nigeria
[8]McGill University, Montréal, Canada
[9]Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria

funperefagha100@gmail.com

Abstract. Recent advances in open-source magnetic resonance imaging (MRI) hardware, have significantly improved engineering of low field (LF) magnetic strengths, making LF viable for a broad range of clinical applications. It is regarded as a feasible imaging modality in low and middle-income countries (LMICs), where imaging resources remain highly constrained. However, LF MRI remains susceptible to noise, artifacts, coil variations, and acceleration-related degradation. These challenges raise critical questions as to whether the images acquired from these scanners can be trusted for safe clinical diagnosis and decision-making. To address this, a unified epistemic and aleatoric uncertainty enhanced autoencoder framework was used to analyse a scan acceptance rule using the open source OSI2 ONE k-space dataset from Leiden University Medical Center (LUMC). The result motivated the development of a two-axis uncertainty-based scan acceptance framework with separate percentile-derived thresholds. The framework successfully identified 42.3% of LUMC, unsafe low-rank reconstruction (LLR) scans while maintaining a zero false-rejection rate for safe scans.

Keywords: Low-field MRI, Uncertainty, Autoencoder.

1 Introduction

Low-field (LF) MRI which operates at magnetic field strengths below 1T has emerged as an affordable alternative to High-field MRI [2] offering several advantages such as portability, reduced system cost, improved safety profile, and the ease to acquire scans in low-resource settings [11]. It is gaining ground as a way to accelerate the accessibility of MRI machines in low-resource environments. Despite its lower magnetic field strength, it has demonstrated significant diagnostic performance comparable to high field MRI for neurological imaging [11]. However, the images acquired from LF MRI remain vulnerable to noise, artifacts, coil variations, acceleration-related degradation, low signal-to-noise ratio (SNR) and longer scan times. To address these challenges, accelerated image reconstruction techniques are used to reduce scan time during data acquisition by undersampling k-space. Recent advances in open-source MRI hardware and computational reconstruction techniques have significantly improved LF-MR image quality, making them viable for a broad range of clinical application [21]. Yet, there is still a significant gap in interpretation of its images in a low-resourced setting.

Deep learning (DL) methods in MRI reconstruction are currently growing rapidly offering faster reconstruction times and mitigating artifacts and noise than iterative methods [17]. However, they are not without limitations. They may generate false features (hallucinations) when applied to a data distribution not seen during training and any small distributional shift in the data can cause the model to be unstable [16]. Furthermore, an undersampled k-space can equally make DL models inherent to uncertainty [4], raising the crucial need to estimate and quantify uncertainty of DL approaches [19]. Uncertainty in DL is broadly grouped into aleatoric and epistemic. Aleatoric uncertainty arises from noise in the data and cannot be reduced by acquiring more data whereas epistemic uncertainty reveals the model's limitation in understanding the data and can be reduced by additional training data [10]. Most existing works have explored uncertainty in DL-based MRI reconstruction models [4,15,3] meanwhile very few works focus on quantifying both forms of uncertainty in MRI reconstruction particularly in LF-MRI.

To address this, we propose a unified uncertainty-aware reconstruction network built using variational autoencoder (VAE) with hybridized aleatoric and epistemic uncertainty quantification. The output from the proposed algorithm is further used to create a triage acceptance rule, classifying each LF scan as accept-able, uncertain or repeat acquisition. The experiments in this study were performed using the open-source OSI^2 ONE LF-MRI (47mT) k-space dataset [21]

2 Related Works

Several studies have incorporated uncertainty estimation into MRI reconstruction. Aleatoric uncertainty has been explored using model-based DL (MoDL) [22], variational autoencoders with Monte-Carlo sampling and a Stein's Unbiased Risk Estimator (SURE) [3], conformal prediction methods [6], and pixel classification estimation approach [4] which shows enhanced reconstruction reliability and calibrated uncertainty.

Epistemic Uncertainty has been addressed through Bayesian variational networks [15] and diffusion-based approaches [5] that model reconstruction as a probabilistic inverse problem and estimate uncertainty from posterior distribution. Other works jointly models both uncertainties using Monte Carlo dropout [18], deep ensembles [13] and Bayesian frameworks [12].

In low-field setting, Anyimadu et al. [1] estimated uncertainty using a deep ensemble approach through cross-validation, though the model was trained using simulated low-field MRI data derived from the Human Connectome Project (HCP) dataset. More recently, Zhang et al. [23] introduced ReDiff, a reliability aware diffusion framework that utilizes spatial uncertainty maps to reduce hallucinations during low-to-high-field MRI synthesis. Despite the efficiency of existing reconstruction models with uncertainty quantification, focus is primarily on uncertainty estimation accuracy rather than downstream clinical decision support, particularly in low-field MRI in resource constrained settings where non-specialist operators require actionable scan quality guidance.

3 Methods

3.1 Data Description and Preprocessing

For this study, we utilized the publicly available low-field MRI dataset (OSI^2 ONE) which was acquired using a 47mT Halbach-based MRI system constructed at the Leiden University Medical Center (LUMC) [21]. It comprises of five different subsets of healthy brain MRI scans which were acquired with the intention of evaluating undersampling strategies, acquisition designs and reconstruction methods tailored for LF imaging. The subsets include, Partial Fourier (PF; n = 16), RF Coil Comparison (n = 70), Locally Low-Rank (LLR) Reconstruction (n = 52), LORAKS Reconstruction, and Repeatability (n = 75).

The MRI data were acquired in the ISMRM Raw Data Format (ISMRMRD) [7]. Using HDF5 utilities, the repeatability and standard k-space acquisitions were reconstructed into paired lists of degraded inputs and clean target slices. These pairs were transformed into PyTorch tensors. The repeatability k-space data were split into training (80%) and validation (20%) sets. The training split was made at slice level. Finally, data from five under-sampling strategies experimented at LUMC were reconstructed to serve as the testing dataset for model evaluation.

3.2 Low Field MRI Scan Acceptance Rule

Visual Quality control (QC) in medical imaging is the process of establishing whether a scan or dataset meets a required set of standards [14]. The reconstructed low-field MRI are affected by acquisition noise, resolution, and/or image artefacts, and traditional standard for identification of these issues are labour-intensive manual visual inspection. However, uncertainty has been identified to be of great importance for decision making, and it is anticipated to increase patient safety in reduction in harm from mis- or no-diagnosis. It is projected to allow for robust rejection of unreliable predictions or out-of-distribution samples.

In this study, the two categories of uncertainty quantification were hybridized to analyse the acceptance rule for VAE reconstructed low-field MRI. The proposed rule will be adapted to follow a 4-quadrant Scan Acceptance Matrix [20] depicted in Figure 1. It is also referred to as a downstream deployment framework used to identify sources of uncertainty that will support the management of clinical logistics as shown in the chart. Meanwhile, the effective low and high thresholds for both uncertainty categories have to be pre-determined, since they are measure in different units.

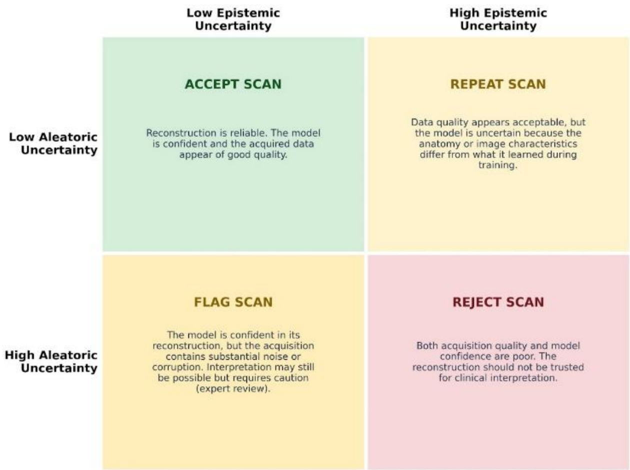

Fig. 1. A 4-quadrant Scan Acceptance Matrix [20]

3.3 Model Architecture

The reconstruction algorithm is a probabilistic deep learning model designed for both image reconstruction and uncertainty quantification in low-field MRI. The architecture following the Variational Autoencoder (VAE) [9] was configured with task-specific inductive biases on accelerated k-space reconstruction. The model takes an inverse fast Fourier transform (iFFT) reconstructed image as input and outputs a mean reconstruction. A key architectural contribution in this design is the dual-head design shown in Figure 2, which also produces two independent outputs. The mean head is the estimated pixel-wise expected reconstruction ˆμ, fine-tuned to minimize the composite reconstruction loss, and the log-variance head measures the pixel-wise aleatoric (data-inherent) uncertainty $\log(\sigma^2)$**.** The exponentiated output, $elog(\sigma^2)$, constitutes the aleatoric uncertainty map $U_{aleatoric}$). The MC Dropout remains active throughout the decoder at inference time. Running $T = 20$ stochastic forward passes through the full encoder–decoder pipeline and computing the pixel-wise variance across the T mean reconstructions yields the epistemic uncertainty map $U_{epistemic}$) - the component attributable to model uncertainty rather than data noise. The concatenation of both uncertainties is defined in Equation 1.

$$U_{total} = U_{epistemic} + U_{aleatoric} \quad (1)$$

Model Training: Training was conducted in PyTorch on an NVIDIA Tesla T4 GPU for up to 200 epochs, with early stopping triggered after 30 epochs without validation loss improvement. The best checkpoint was saved at epoch 111 of 141 total epochs. For the model optimization, we used Adam optimizer with a learning rate of 1×10^{-4} with a batch size of 8.

$$\mathcal{L} = \mathcal{L}_{L1} + \mathcal{L}_{SSIM} + \beta \, . \mathcal{L}_{KL} \quad (2)$$

Where $\mathcal{L}_{L1}$ is the mean absolute pixel error, $\mathcal{L}_{SSIM}$ is the Structural Similarity Index loss, and LKL is the KL divergence term with an annealing schedule, weighted at $\beta = 1 \times 10^{-4}$.

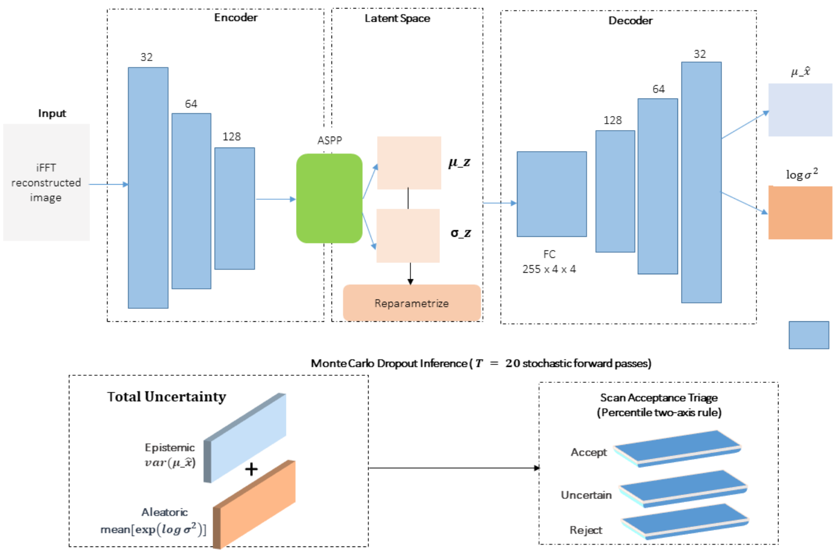

Fig. 2. Diagram of the Proposed Model Architecture

4 Results

As shown in Table 1, the model achieved a best validation loss of 0.1794, with structural similarity index (SSIM) of 0.3926, and normalized mean squared error (NMSE) of 0.18225, each marginally out performing corresponding training-set metrics, indicating the absence of over fitting. The training history confirmed stable, monotonic convergence across all four tracked metrics, with no evidence of loss divergence or instability in the KL-annealing schedule. To assess generalisation beyond the fixed 80/20 split, $k$fold cross-validation was performed at $k \in \{2,4,8\}$ across a pooled set of 3,000 training slices. The results as summarised in Table 1, shows the performance improved marginally and consistently as k increases. SSIM increased from 0.3498 at 2-fold to 0.3604 at 8-fold, and NMSE decreased from 0.17676 to 0.16916. Standard deviations remained small across all folds, confirming that the architecture generalises consistently and is not overfit to any particular data partition. However, the model at 0-fold retain its' position as improvement on model attained at 8-folds cross-validation (Table 1).

Table 1. Cross-validation across 2/4/8 folds (3,000 pooled training slices) confirms the architecture generalises consistently and was not overfit to the original 80/20 split.

Fold	Recons. Error (Loss)	PSNR	SSIM	NMSE	L1	SNR	CNR
0-fold	0.1794	18.86	0.3926	0.18103	0.0750	15.47	13.33
2-fold	0.1953 ± 0.0009	18.33± 0.03	0.3498± 0.0030	0.17676 ± 0.00093	0.0837 ± 0.0003	11.04 ± 0.01	9.02± 0.00
4-fold	0.1928± 0.0012	18.41± 0.07	0.3571± 0.0034	0.17165 ± 0.00182	0.0826 ± 0.0007	11.83 ± 0.27	9.73± 0.22
8-fold	0.1915± 0.0017	18.43± 0.05	0.3604± 0.0071	0.16916 ± 0.00178	0.0818 ± 0.0007	11.92 ± 0.72	9.75± 0.68

The 0-folds cross-validated model was evaluated across the OSI² ONE dataset four acquisition conditions: Fully Sampled, Accelerated ($R = 2$), Partial Fourier, and LLR/Under sampled, and their mean normalized uncertainty were compared across acquisitions. From Figure 3, Partial Fourier scans had the lowest mean U_{norm} (0.2751), reflecting more stable reconstruction quality under this acquisition strategy, while Accelerated and LLR show roughly equal levels of highest mean uncertainty. A Kruskal-Wallis test across all four conditions yielded $H = 6.51$, $p = 0.0891$, indicating no statistically significant overall group difference at the $\alpha = 0.05$ threshold, while the LLR / Under sampled has the widest spread, stretching from 0.0 to 1.0, meaning it is highly unpredictable.

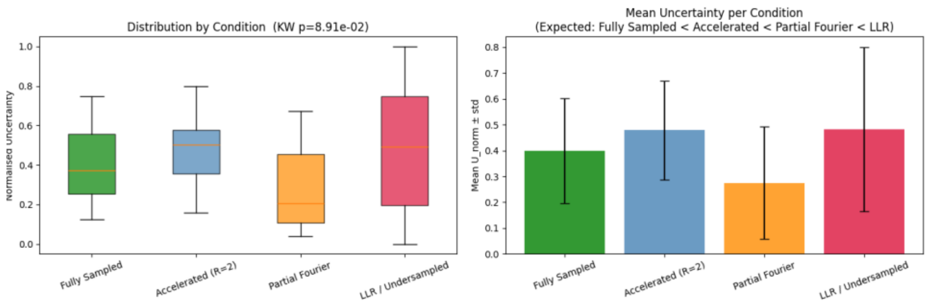

Fig. 3. Distribution of normalised uncertainty across the four acquisition conditions

Furthermore, Spearman rank correlations (ρ) were computed between the three uncertainty estimates ($U_{aleatoric}$, $U_{epistemic}$), and SSIM across acquisition conditions and stratified per condition. Results are presented in Table 2.

Table 2. Uncertainty Pattern across Reconstruction Quality

Condition	Epistemic (ρ)	Aleatoric (ρ)	Volume(n)
Fully Sampled (R=1)	-0.8741	0.1538	12
Accelerated (R=2)	-0.6993	0.0350	12
LLR undersampled (R=1.7)	-0.4286	-0.3571	7

A correlation analysis was performed (Table 2) around the low quality accelerated conditions ($R = 2$) under-sampled at 50% k-space, the LLR ($R = 1.7$) images under sampled at 70% of the k-space and the fully sampled ($R = 1$) k-space images. It was discovered that epistemic uncertainty was significantly and negatively correlated with SSIM ($\rho = -0.8741$), and the next ranked correlation is the accelerated under samples at SSIM correlation of ($\rho = -0.6993$), while the LLR undersampled attained a weaker correlation with SSIM ($\rho = -0.4286$), which was assumed to have been as a result of the less number of data samples ($n = 7$) in LLR undersampled. However, the consistent negative epistemic uncertainty and SSIM across the different conditions suggests that higher epistemic uncertainty is associated with lower reconstruction quality. In contrast, aleatoric uncertainty showed weak and unstable correlation with SSIM, changing signs across acquisitions conditions.

In addition, the stability analysis of the repeatability scan shows low coefficients of variation (~ 12%), indicating that uncertainty estimates are reproducible across acquisition settings and across repeat sessions. Therefore, thresholds for each uncertainty axis were derived independently from the validation set distribution using robust percentile clipping. Raw epistemic and aleatoric uncertainty values were first normalized to [0,1] by clipping at the 2nd and 98th percentiles of their respective distributions, suppressing the influence of outlier slices. The low/high boundary for each axis (τlow) was then set at the 50th percentile (median) of the normalized distribution, and the upper boundary ($\tau high$) at the 75th percentile, yielding data-driven thresholds that reflect the natural operating range of each signal independently.

$$Epistemic\ Threshold = \begin{cases} H_{epistemic}\ \tau_{high} \geq P_{75}\ (U_{epistemic}) \\ L_{epistemic}\ \tau_{low} < P_{50}\ (U_{epistemic}) \end{cases} \tag{3}$$

$$Aleatoric\ Threshold = \begin{cases} H_{aleatoric}\ \tau_{high} \geq P_{75}\ (U_{aleatoric}) \\ L_{aleatoric}\ \tau_{low} < P_{50}\ (U_{aleatoric}) \end{cases} \tag{4}$$

Where $H_{epistemic}$ and $H_{aleatoric}$ are the high epistemic and aleatoric values, and $L_{epistemic}$ and $L_{aleatoric}$ are the low epistemic and aleatoric values. Here, $U_{epistemic}$ (percentile clipped) epistemic and aleatoric uncertainty signals respectively, and P50, P75 denote the 50th and 75th percentiles of the validation set distribution (see Equations 3, 4). The low epistemic is below 0.5367 threshold, while a value above 0.7718 represents high epistemic uncertainty. A value below 0.5119 represents low aleatoric in the data, while a value above 0.8029 represents highly aleatoric.

The safety evaluations were performed using acquisition conditions as reference indicators rather than clinically assessed labels. The evaluation scans comprised 92 scans (52 LLR/undersampled, 24 fully sampled or accelerated, 16 partial Fourier) of the OSI[2] dataset. When all 52 LLR scans were processed by the two-axis rule (see Figure 4) enhanced VAE, the framework identified 22 of 52 scans for repeat acquisition (sensitivity = 42.3%), with zero fully sampled or accelerated scans incorrectly rejected (false reject rate = 0.0%), yielding a Youden index of 0.423.

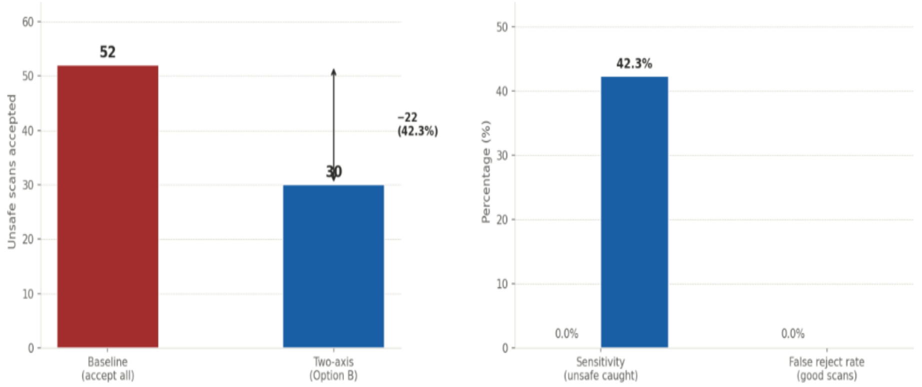

Fig. 4. Representation of Safety Evaluations of LLR Unsafe Scans with hybridized uncertainty aware VAE

5 Discussions

This study presents a hybridized epistemic–aleatoric uncertainty-aware autoencoder for scan-quality assessment in low-field MRI. Among the two uncertainty measures, epistemic uncertainty emerged as the primary, reliable driver of reconstruction-quality assessment. Reconstruction quality remained modest, with a best validation SSIM of approximately 0.39. Nevertheless, cross-validation demonstrated consistent performance across data splits, indicating that the model generalized reasonably well despite the limited size of the dataset. The relatively low reconstruction metric attained may be partly attributed to the low quality of the input data and the VAE's tendency to produce overly smooth reconstructions, which can result in the loss of fine image details. A key finding of this study was the strong inverse relationship between epistemic uncertainty and reconstruction quality. Epistemic uncertainty was significantly negatively correlated with SSIM across acquisition conditions, demonstrating that model confidence decreased as reconstruction difficulty increased. This suggests that epistemic uncertainty provides a meaningful estimate of reconstruction reliability and may serve as a useful indicator of scans requiring additional review. In contrast, aleatoric uncertainty showed inconsistent correlations with SSIM and exhibited poor discriminative performance. The near constant behavior of the aleatoric uncertainty estimates indicates that the variance prediction branch primarily captured baseline image noise rather than clinically relevant reconstruction degradation. This is consistent with the aleatoric dominance effect reported by Kendall and Gal [8], who found that the log-variance head tends to capture residual data noise rather than true reconstruction degradation; nevertheless, this does not undermine the prospect of combining both uncertainty types for the management of clinical logistics.

The repeatability analysis further demonstrated that epistemic uncertainty was reproducible across repeated acquisitions while remaining responsive to acquisition-related changes. Although aleatoric uncertainty showed even higher repeatability, its minimal variation across conditions suggests limited utility as a standalone marker of scan quality. Following the recommendation of Kendall and Gal [8] to treat epistemic and aleatoric uncertainty as fundamentally distinct signals requiring separate decision boundaries, independent percentile-derived thresholds were established for each axis to construct a two-axis scan acceptance framework. Applied to the dataset, this rule identified 42.3% unsafe LLR-based undersampled scans while maintaining a zero false rejection rate on safe scans (fully sampled and accelerated acquisitions). While the 42.3% detection rate remains modest, the ground truth labels in this preliminary evaluation were defined by acquisitions conditions of image quality rather than independent clinical measures of scan acceptability. The results suggest that uncertainty estimates may support practical scan-quality triage by identifying acquisitions that are safe or requires reacquisition. Future studies will validate this framework against clinical evaluations and semantic image quality metrics, and compare with state-of-art reconstruction models and uncertainty quantification methods.

6 Impact in RCS

Overall, the findings suggest that hybridized uncertainty-aware reconstruction may provide a useful approach for automated scan-quality triage in low-field MRI. Low-field MRI is susceptible to challenges such as noise, artifacts, coil variation, and acceleration related degradation which may also be relevant in resource constrained settings (RCS) including Sub-Saharan Africa (SSA), where scanners frequently exhibit poorer contrast and resolution under hardware and operational constraints. Although LUMC dataset used in this study cannot fully represent specific imaging conditions or patient condition found in SSA, its variations provides a useful basis for examining how uncertainty changes under different image quality conditions. Therefore, this study is relevant to real world LF-MRI settings, where the safety status (accept, reject, flag or repeated) of scans is critical for medical reporting. In addition to reducing the workload of the limited number of MRI physicists in RCS, an AI-enhanced scan acceptance rule is expected to lower the expertise barrier required to operate low-field MRI scanners in low-resource clinics. Future work should investigate higher-performing reconstruction architectures, larger training datasets, including low-field MRI data from SSA and other resource constrained settings as well as improved uncertainty calibration strategies to further enhance triage sensitivity and clinical utility.

Acknowledgements. This work was part of the Sprint AI Training for African Medical Imaging Knowledge Translation (SPARK) Academy 2026 summer school on deep learning in medical imaging. The authors would like to thank the instructors of the summer for providing insightful background knowledge on brain tumours that informed the research presented here, most notably: Maruf Adewole, Mohannad Barakat, Craig Jones, Noha Magdy, Amal Saleh, Aondona Iorumbur, Bernes Atabonfack, Confidence Raymond, Craig Jones, Ilerioluwakiiye Abolade, Jeremiah Fadugba, Juampablo Heras Rivera, Kenneth Aguh, Lukman E. Ismaila, Mohtady Barakat, Nourou Dine Bankole, Saikat Roy, Toufiq Musah, Willem Boonzaier.. The authors acknowledge the funding support provided to SPARK through the Lacuna Fund for Health and Equity (PI: Udunna Anazodo), the RSNAR & E Foundation (PI: Farouk Dako), the University of Washington Population Health Initiative Tier2 Grant (PI: Mehmet Kurt), McGill University Healthy Brain and Healthy Lives (HBHL, Anazodo), and the National Science and Engineering Research Council of Canada (NSERC) Discovery Launch Supplement (PI:UdunnaAnazodo, DGECR-2022-00136).

References

1. Anyimadu, D.T., Abdelsamea, M.M., Eldaly, A.K.: Direct low field MRI super-resolution using undersampled k-space (Feb 2026). https://doi.org/10.48550/arXiv.2603.00668Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016)
2. Arnold, T.C., Freeman, C.W., Litt, B., Stein, J.M.: Low-field MRI: Clinical promise and challenges. Journal of Magnetic Resonance Imaging 57(1), 25–44 (Jan 2023). https://doi.org/10.1002/jmri.28408Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
3. Edupuganti, V., Mardani, M., Vasanawala, S., Pauly, J.: Uncertainty Quantifica tion in Deep MRI Reconstruction. IEEE Transactions on Medical Imaging 40(1), 239–250 (Jan 2021). https://doi.org/10.1109/TMI.2020.3025065
4. Ekanayake, M., Pawar, K., Chen, Z., Egan, G., Chen, Z.: PixCUE: Joint Uncer tainty Estimation and Image Reconstruction in MRI using Deep Pixel Classifica tion. Journal of Imaging Informatics in Medicine 38(4), 2071–2084 (Dec 2024). https://doi.org/10.1007/s10278-024-01250-3
5. Eldaly, A.K., Figini, M., Alexander, D.C.: Bayesian Uncertainty-Aware MRI Re construction (Mar 2026). https://doi.org/10.48550/arXiv.2603.13439
6. Giannakopoulos, I.I., Muthukumar, L.B.G., Lui, Y.W., Lattanzi, R.: Pixel wise Uncertainty Quantification of Accelerated MRI Reconstruction (May 2026). https://doi.org/10.48550/arXiv.2601.13236
7. Inati, S.J., Naegele, J.D., Zwart, N.R., Roopchansingh, V., Lizak, M.J., Hansen, D.C., Liu, C.Y., Atkinson, D., Kellman, P., Kozerke, S., Xue, H., Campbell Washburn, A.E., Sørensen, T.S., Hansen, M.S.: ISMRM Raw data format: A proposed standard for MRI raw datasets. Magnetic Resonance in Medicine 77(1), 411–421 (Jan 2017). https://doi.org/10.1002/mrm.26089
8. Kendall, A., Gal, Y.: What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? (Oct 2017). https://doi.org/10.48550/arXiv.1703.04977
9. Kingma, D.P., Welling, M.: Auto-Encoding Variational Bayes (Dec 2022). https://doi.org/10.48550/arXiv.1312.6114
10. Kiureghian, A.D., Ditlevsen, O.: Aleatory or epistemic? Does it matter? Structural Safety 31(2), 105–112 (Mar 2009). https://doi.org/10.1016/j.strusafe.2008.06.020

11. Kravchenko, D., Hagar, M.T., Vecsey-Nagy, M., Kabat, I., Groteklaes, A., Luetkens, J.A., Kuetting, D., Isaak, A., Emrich, T., Varga-Szemes, A., Spampinato, M.V.: Low-field and portable MRI technology: Advancements and innovations. European Radiology Experimental 9(1), 103 (Oct 2025). https://doi.org/10.1186/s41747-025-00638-2
12. Kuroki, S., Fujita, N., Terada, Y.: A unified approach for maintain ing MRI reconstruction quality and quantifying both aleatoric and epis temic uncertainty. Magnetic Resonance Imaging 131, 110680 (2026). https://doi.org/10.1016/j.mri.2026.110680
13. Küstner, T., Hammernik, K., Rueckert, D., Hepp, T., Gatidis, S.: Predictive un certainty in deep learning–based MR image reconstruction using deep ensembles: Evaluation on the fastMRI data set. Magnetic Resonance in Medicine 92(1), 289–302 (Jul 2024). https://doi.org/10.1002/mrm.30030
14. Laves, M.H., Ihler, S., Fast, J.F., Kahrs, L.A., Ortmaier, T.: Recalibration of Aleatoric and Epistemic Regression Uncertainty in Medical Imaging (Apr 2021). https://doi.org/10.48550/arXiv.2104.12376
15. Luo, G., Blumenthal, M., Heide, M., Uecker, M.: Bayesian MRI Reconstruction with Joint Uncertainty Estimation using Diffusion Models. Magnetic Resonance in Medicine 90(1), 295–311 (Jul 2023). https://doi.org/10.1002/mrm.29624
16. Muckley, M.J., Riemenschneider, B., Radmanesh, A., Kim, S., Jeong, G., Ko, J., Jun, Y., Shin, H., Hwang, D., Mostapha, M., Arberet, S., Nickel, D., Ramzi, Z., Ciuciu, P., Starck, J.L., Teuwen, J., Karkalousos, D., Zhang, C., Sriram, A., Huang, Z., Yakubova, N., Lui, Y.W., Knoll, F.: Results of the 2020 fastMRI Challenge for Machine Learning MR Image Reconstruction. IEEE Transactions on Medical Imaging 40(9), 2306–2317 (Sep 2021). https://doi.org/10.1109/TMI.2021.3075856
17. Safari, M., Eidex, Z., Chang, C.W., Qiu, R.L.J., Yang, X.: Advancing MRI Re construction: A Systematic Review of Deep Learning and Compressed Sensing Integration. Biomedical Signal Processing and Control 111, 108291 (Jan 2026). https://doi.org/10.1016/j.bspc.2025.108291
18. Schlemper, J., Castro, D.C., Bai, W., Qin, C., Oktay, O., Duan, J., Price, A.N., Hajnal, J., Rueckert, D.: Bayesian Deep Learning for Accelerated MR Image Re construction. In: Knoll, F., Maier, A., Rueckert, D. (eds.) Machine Learning for Medical Image Reconstruction, vol. 11074, pp. 64–71. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-00129-2_8
19. Seoni, S., Jahmunah, V., Salvi, M., Barua, P.D., Molinari, F., Acharya, U.R.: Application of uncertainty quantification to artificial intelligence in healthcare: A review of last decade (2013–2023). Computers in Biology and Medicine 165, 107441 (Oct 2023). https://doi.org/10.1016/j.compbiomed.2023.107441
20. Singh, Y., Andersen, J.B., Hathaway, Q., Venkatesh, S.K., Gores, G.J., Erickson, B.: Deep learning-based uncertainty quantification for quality assurance in hepatobiliary imaging-based techniques. Oncotarget. 2025 Apr 4. https://doi.org/10.18632/oncotarget.28709
21. van den Broek, R., Lena, B., Najac, C., Kolbitsch, C., Dong, Y., Ilicak, E., Imre, B., Staring, M., Webb, A.: Low-field MRI dataset with OSI^2 ONE [Dataset] (Apr 2026). https://doi.org/10.5281/ZENODO.19661402
22. Wang, K., Angelopoulos, A., Alfredo, D.G., Amit, K., Efrat, S., Stella, Y., Jitendra, M., Michael, L.: Rigorous Uncertainty Estimation for MRI Reconstruction (2022)
23. Zhang, Z., Jing, P., Yuan, R., Hu, L., Wang, A., Wang, F., Wu, Y., Islam, K.T., Chen, Z., Wang, Z., Lally, P., Yang, G.: Towards Trustworthy Selective Generation: Reliability-Guided Diffusion for Ultra-Low-Field to High-Field MRI Synthesis (Mar 2026). https://doi.org/10.48550/arXiv.2603.11325